

ASSESSMENT OF CONSONANT ARTICULATION IN GLOSSECTOMEE SPEECH BY DYNAMIC MRI

K. Mády¹, R. Sader¹, A. Zimmermann¹, P. Hoole², A. Beer³, H.-F. Zeilhofer¹, Ch. Hannig³

¹ Department of Oral and Maxillofacial Surgery, University of Technology, Munich

² Department of Phonetics and Linguistic Communication, Ludwig Maximilian University, Munich

³ Department of Radiology, University of Technology, Munich

mady@mkg.med.tu-muenchen.de

ABSTRACT

For the speech ability assessment of glossectomee patients, dynamic MRT was used in order to obtain real time imaging of articulatory movements. At the present state, the method is best suited for continuous consonants, but it can be used for the evaluation of any speech sound. In this study, pre- and postoperative consonant articulation of eight patients with cancer of the oral cavity is compared by means of acoustic analysis and real-time MRI.

1. INTRODUCTION

It is nearly a commonplace that a general description of oral and mandibular muscle function alone does not allow for exact predictions regarding articulatory ability after tongue surgery. Not only the tongue reconstruction method, but also compensatory skills and personal motivation have an influence on the severity of postoperative articulatory impairments. Several acoustic studies have been done lately in order to evaluate post-operative speech ability of patients who have undergone a resection and reconstruction of the tongue and/or floor of the mouth ([1], [2]). In the case of impaired speech, real-time MRI is a promising means to a better understanding of articulatory and compensatory processes. In the present paper, first results of a long-term study will be presented.

2. MATERIAL AND METHODS

The results to be presented in this study are based on eight patients with a cancer of the anterior and lateral oral cavity, the tongue and/or the floor of mouth being concerned. They had no speech or language disorders before their current disease, nor any tumours in the head and neck area previously. The recordings were made a few days before the operation and approximately four weeks postoperatively, i. e. before an optional radiotherapy would begin.

The test material contained existing German words with simple CV-structures. Real words proved to be the only possible test material, since the communication environment during the MRI recordings would make understanding of

nonwords very difficult. Apart from the MRI sessions, "clean" acoustic recordings were made in a silent environment. In addition to the test words, extreme tongue movements, like protruding the tongue, moving it up and down to the maximum extension, or pulling it backwards towards the velum, were recorded via MRI. As oral cancer often causes a swallowing disability, the recordings were completed by a swallow test.

At the present state of our study, the possible maximum recording speed is 8 images per second, using a T1 fast gradient echo sequence on Philips ACS NT Gyroscan (TR = 4.0 ms, TE = 1.1 ms, fa = 6, slice thickness = 10 mm) with the new "sensitivity encoding" system (SENSE). With this feature combination the entire head and neck area can be viewed, while resolution and slice thickness are sufficient for gaining information about the movement of all articulators from the lips to the larynx, and also, the contrast between tissue and air is sharp enough. In addition, images that are computed in the SENSE system lack the artefacts (reflection of the back part of the head to the front) that occurred in the previous gradient echo recordings.

Objective speech assessment was obtained from the Munich Intelligibility Profile (MIP) ([3]) that gives information about impaired (= incorrectly identified) phonological features for each patient, relying on the identification of words by listeners who do not know the patients nor are trained in speech diagnostics.

In order to achieve a better understanding of the connection of muscle function and speech ability, the resection and reconstruction of the defect were documented in a surgical mapping protocol ([1]).

3. RESULTS

3.1. Resection and reconstruction

The resection concerned the floor of the mouth and the tongue base in five cases (Patients 2, 4, 5, 9, and 13) and included the tongue surface in three cases (Patients 1, 3, and 15). Size and localisation of the defect are shown in Fig. 1. In each case, an immediate reconstruction was done, where the defect was closed primarily (by fixing the tongue to the floor of the mouth) or by a local flap. All patients showed a restricted tongue mobility after the operation.

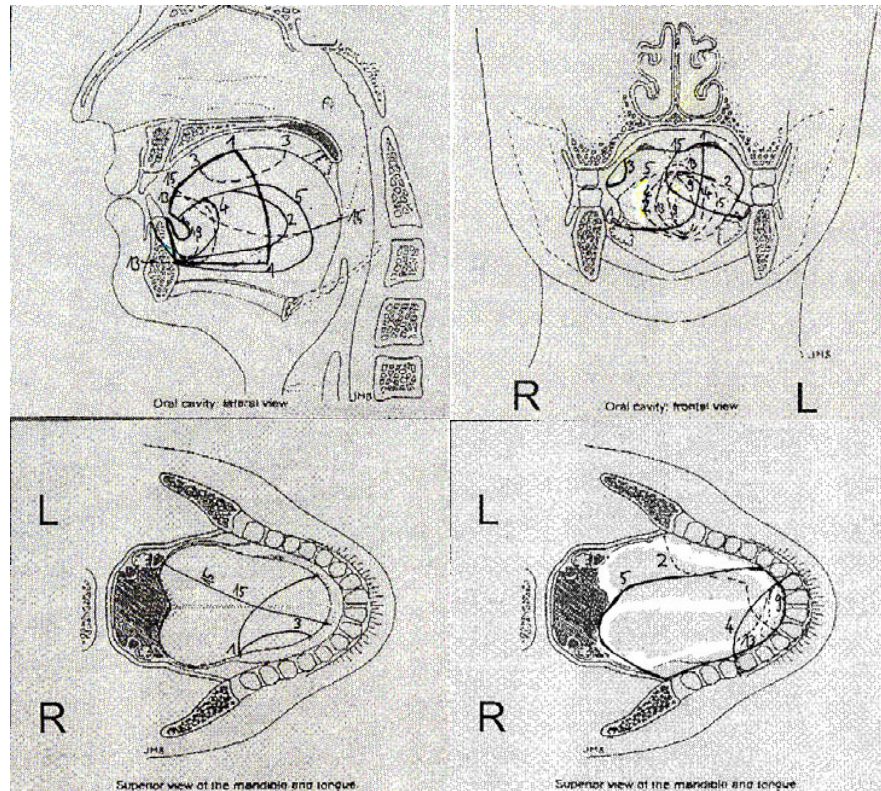


Figure 1: Surgical Mappings of the patients, (a) lateral view, (b) frontal view, (c) superior view with tongue participation, and (d) superior view without tongue participation.

3.2 Speech ability

Patients 3, 4, 9, and 13 show no or minimal articulatory impairment, whereas Patients 2, 1, 15, and 5 show articulatory disabilities ranging from just noticeable up to very severe. This finding correlates roughly with the tumour size in these patients. Concerning the judgements made in MIP, no sound was systematically “misheard” by listeners for the first group. For the more severely disabled group, both place and manner of articulation were identified incorrectly. For Patient 1 there is a tendency to mishear his velar sounds as anterior (alveolar and labial) sounds, whereas for Patient 2 the misidentifications mainly regarded stops that were perceived as fricatives. No clear tendency could be observed for Patients 15 and 5 who both had a severe articulatory impairment.

3.3 Assessment by dynamic MRI

In order to exemplify the possibilities given by real-time MRI, we will present a short analysis of articulatory changes regarding two sounds: /z/ and /t/.

3.3.1 /z/

Voice quality may vary widely in German /z/, depending on phonotactic and dialectal circumstances. The sound in the target word *Rosi* would be pronounced [z] in Standard Modern German but is often pronounced completely voiceless in Bavaria where the recordings took place.

No change in the articulation of /z/ was observed for Patients 2, 3, 4, and 9. Despite the tongue fixation, the remaining free tip of the tongue was sufficient for the alveolar articulation.

The postoperative /z/ articulation of Patient 13 is somewhat unclear; it sounds palatalised. This impression is accounted for by the MR imaging: the patient cannot produce an apical-alveolar contact, i. e. the contact surface is larger and shifted towards the hard palate (see Fig. 2). It is interesting that two of three postoperative /z/ sounds are voiced as opposed to three voiceless sounds preoperatively.

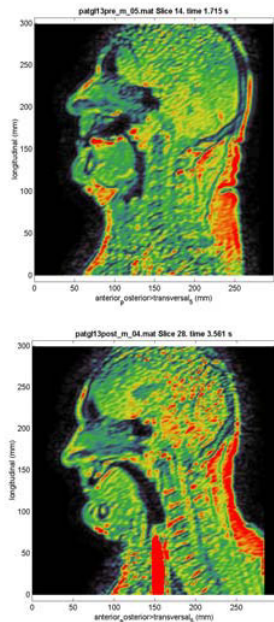


Figure 2: Articulation of /z/ in Germ. *Rosi* by Patient 13 (a) before and (b) after the operation.

Although the speech of Patient 15 is largely distorted, the articulation of /z/ is not affected. It can be explained by the fact that the large amount of missing tongue tissue does not affect the articulation of the unvoiced fricative if enough friction intensity can be produced and the tip of the tongue is capable of an alveolar constriction (Fig. 3).

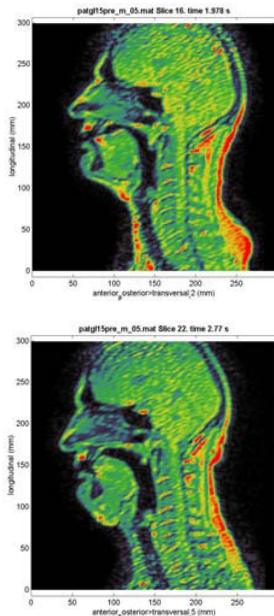


Figure 3: Articulation of /z/ in Germ. *Rosi* by Patient 15 (a) before and (b) after the operation.

Both Patient 1 and 5 show a similar articulation pattern to Patient 13, the constriction surface being larger. Both the auditory and acoustic image (spectrogram) show that the sound is pronounced as an approximant without friction. In case of Patient 5, the sound appears to be a labial approximant – this is also supported by the MR scan (Fig. 4) where the tongue tip touches not only the alveolus but also the upper lip.

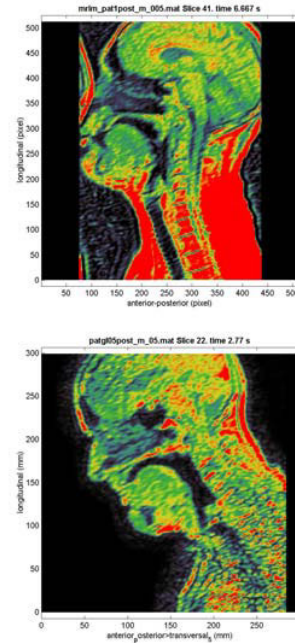


Figure 4: Postoperative MR images of the articulation of /z/; (a) Patient 1, (b) Patient 5.

3.3.2 /r/

The German phoneme /r/ has a wide range of allophonic variation, the uvular trill [R] being the standard pronunciation, whereas among speakers of the Bavarian regional variety, the apical realization [r] is fairly widespread. It can be claimed that all of our patients would have difficulties with producing an apical trill postoperatively, so they might choose to produce a velar or uvular trill, flap or fricative instead. Therefore we looked at the postoperative r-realizations of those patients who used the apical allophone prior to the operation.

An apical realization of /r/ was observed in Patient 1, 2, 5 and 13. Although the tongue mobility does not allow for an apical trill production postoperatively, Patient 1 and 13 stick to the alveolar realization and produce an approximant-like sound. Patient 2 shifts to a uvular/pharyngeal articulation. Patient 5 seems to keep the same tongue form he had for the apical trill, but the exact place of articulation is hard to tell because of the extremely restricted mouth opening that causes all articulations to seem anterior.

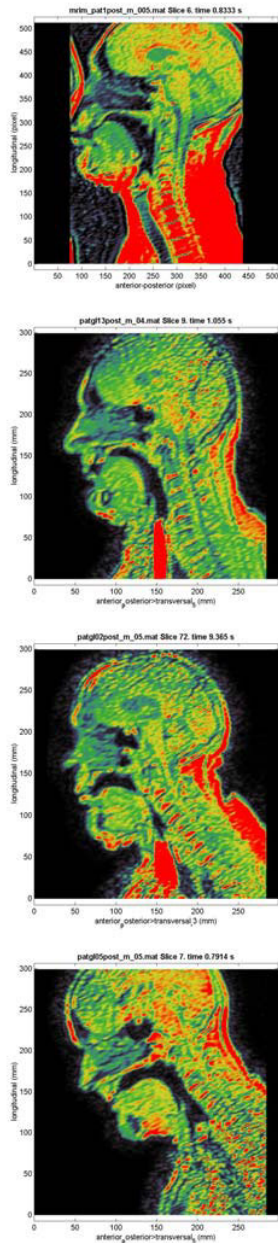


Figure 5: postoperative /r/-realizations of (a) Patient 1, (b) Patient 13, (c) Patient 2, and (d) Patient 5.

4. FURTHER ANALYSIS AND CONCLUSIONS

As mentioned above, the present analysis both of the MR images and of the acoustic recordings is far from being complete. An objective evaluation method of articulatory movements concerning size and location of the constriction, tongue position, lip protrusion and opening is under testing at present. The distance measurements are based on two fixed points: the inferior apex of the alveolar ridge and the anterior inferior edge of the second cervical vertebra (axis). From the midpoint of this line, a grid of 13 vectors was constructed (where adjacent lines are at an angle of 15° to each other). The resulting semicircle covers the entire oral area where consonant articulation takes place. Along each line, two points are measured for every tongue position, beginning from the origin of the grid: (a) the outer edge of the tongue and (b) the closest point of the palate, thus the origin–palate distance is identical with the maximal possible tongue extension (100%) along the line. In compensatory articulation, the lips are likely to participate in some movements where they are normally not involved. Therefore the lip distance is measured separately, bilabial distance and lip protrusion being the two main factors. The size and location of the constriction is calculated for each repetition of the target sounds (3–7 times during 10 sec recording time) so that mean values can be computed for all sounds.

The frequency of 8 images/sec allows for an analysis of continuous consonants like fricatives, and also the articulation of laterals, nasals and trills can be observed sufficiently. However, it is difficult to get reliable images of stop production by this relatively low sample rate. Until the technical development of MR imaging achieves a real movie speed of 25 images/sec, we have to rely on the articulatory tendencies that can be seen when the images are presented as a “slide show”. Even with this less sophisticated method, we can get a much better insight into soft tissue movements than was possible by any other visual recording method (dynamic ultrasound or video fluoroscopy).

4. REFERENCES

- [1] Mackenzie Beck, Janet et al., “Surgical mapping and phonetic analysis in intra-oral cancer”, in W. Ziegler & K. Deger (Eds.), *Clinical Phonetics and Linguistics*, London: Whurr, 481–492, 1998.
- [2] Savariaux, Christophe et al., “Speech production after glossectomy and reconstructive lingual surgery: a longitudinal study”, *Proc. 2nd Int. Workshop Models Analysis Vocal Emissions for Biomedical Applications (MAVEBA) Firenze, Italy*, 2001.
- [3] Ziegler, Wolfram & Hartmann, E., “Das Münchner Verständlichkeitsprofil (MVP): Untersuchungen zur Reliabilität und Validität“, *Nervenarzt*, 64, 653–658, 1993.